DOCUMENT RESUME

ED 193 340 TM 800 661

AUTHOR Huberty, Carl J.; Smith, Jerry D.
TITLE Linear Classification Functions.
PUB DATE Apr 80
NOTE 23p.; Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Analysis of Covariance; *Classification; *Correlation; Criteria; *Discriminant Analysis; Experimental Groups; Mathematical Formulas; *Measurement Techniques; *Multivariate Analysis; Predictive Measurement
IDENTIFIERS *Linear Classification Function; Standard Scores

ABSTRACT

Linear classification functions (LCFs) arise in a predictive discriminant analysis for the purpose of classifying experimental units into criterion groups. The relative contribution of the response variables to classification accuracy may be based on LCF-variable correlations for each group. It is proved that, if the raw response measures are "standardized," the ordering of the variables can be accomplished by ordering the means of the standardized measures within each group. A computational procedure is presented for obtaining these means as well as for obtaining the LCF-variable within-groups correlations. These correlations may be used for two purposes: (a) to characterize each group with respect to variable scores, and (b) to determine the relative contribution of the variables to separate group classification accuracy. It is proved that the LCF-variable correlations for a group are proportional to standardized variable means for that group. The proof enables a researcher to quite readily accomplish the two interpretation purposes. Two data sets are used to illustrate the interpretation of LCFs. (Author/RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED193340

Linear Classification Functions

Carl J Huberty Jerry D. Smith

University of Georgia

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. D. Smith

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 800 661

Paper presented at the annual meeting of the American Educational Research Association, April, 1980, Boston.

Abstract

Linear classification functions (LCFs) arise in a predictive discriminant analysis for the purpose of classifying experimental units into criterion groups. The relative contribution of the response variables to classification accuracy may be based on LCF-variable correlations for each group. It is proved that, if the raw response measures are "standardized," the ordering of the variables can be accomplished by ordering the means of the standardized measures within each group. A computational procedure is presented for obtaining these means as well as for obtaining the LCF-variable within-groups correlations. Two data sets are used to illustrate the interpretation of LCFs.

Linear Classification Functions

Techniques of discriminant analysis may be useful for two different research situations calling for a multivariate analysis of data. One situation involves the analysis of effects that appear as a result of hypothesis testing. [In such a context suppose that N experimental units comprise K well-defined "groups," and for each unit there are measures on p response variables. Note that this includes the context of a "two-way lay-out."] The general objective of the analysis is to describe the resulting effects. The specific purposes may be: (1) To identify underlying dimensions associated with the apparent effects and their relationship to group differences, and (2) To determine the relative contributions of the p variables to the apparent effects. Central to the purposes of this analysis are linear discriminant functions (LDFs) which result from an eigenanalysis. [The maximum number of LDFs for each effect is typically the associated number of degrees of freedom.] Interpretations of LDFs to accomplish the above purposes are made by examining LDF graphical plots, and determining the p LDF-variable within-groups correlations and/or determining the standardized (partial) LDF weights. (See Huberty, 1975, pp. 547-562.) An appropriate term for such an analysis is descriptive discriminant analysis.

A second situation involves the setting-up of a decision rule to be used for classifying units into pre-existing groups. The rule to be set-up is based on the available p-tuple measures on each of N units. The rule is then used to predict group membership of a unit whose membership is in question. [Typically such an analysis has been considered only in the context of a "one-way lay-out."] The decision rule to be used depends upon conditions of the data and upon the amount of information to be considered in the decision

(see Huberty, 1975, pp. 565-581). Though often based on group membership probabilities, the commonly used rule may equivalently be based on linear composites of the response variable scores, or linear classification functions (LCFs). An appropriate term for such an analysis is predictive discriminant analysis.

[It should be noted that LDFs, which are sometimes called "canonical discriminant functions," are not the same as LCFs. LDF weights are obtained via an eigenanalysis (see Tatsuoka, 1971, p. 161), while LCF weights are obtained from a transformation of the group means, as will be pointed out below. Green (1979) discusses the relationship between LCF weights and LDF weights.]

Thus, discriminant analysis techniques are pertinent to analyzing data relevant for two different types of research situations. In the first type where a descriptive discriminant analysis is appropriate, that of group comparisons, the response variables play the role of criterion variables. In the second type where a predictive discriminant analysis is appropriate, that of group assignment, the response variables play the role of predictor variables. The lack of distinction in roles of variables plus that between LDFs and LCFs, has caused some confusion in methodological circles.

In the literature there has been an apparent lack of recognition of, and confusion between, the two purposes of performing a "discriminant analysis." For example, Spector (1977, p. 159) states that, "Discriminant analysis is useful for prediction and classification...," which should imply the use of LCFs. In the succeeding paragraph Spector states that, "The purpose of discriminant analysis is to specify the linear combination of p variables (that maximally separates the groups)...," which should imply the use of LDFs. Another example of potential confusion is in the writing of

Weiss (1976). Weiss (p. 335) states that, "Linear discriminant functions are designed to predict...," and makes it clear that what is being referred to are LDFs. Predictive accuracy is mysteriously associated with each LDF when he states that, "the first set of beta weights represents the discriminant equation (i.e., LDF) with ... the highest hit rate," and that successive LDFs do not provide "as high a degree of prediction as the first" (p. 336). [It is recognized that classification can be accomplished using LDF scores rather than response variable scores. In fact, the SPSS subprogram DISCRIMINANT (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975, pp. 434-467) classifies in the reduced or discriminant space. This approach to classification does have its merits (see Huberty, 1975, pp. 573-574; Tatsuoka, 1971, pp. 232-233; Tatsuoka, 1975, p. 272). With this approach a classification score using the LDF scores as input data, such as an LCF score discussed below, must be obtained.] For more confusion between LDFs and LCFs see O'Grady (1978, p. 472).

The problem of interpreting LDFs has received considerable attention in the literature (see Borgen & Seling, 1978; Bray & Maxwell, Note 1; Wilkinson, 1977; Bavry, Note 2). The interpretation of LCFs, however, has been discussed very little (see Sanathanan, 1975); this is the focus of the current paper.

Linear Classification Functions

Numerous decision rules are available for carrying out a predictive discriminant analysis. When certain data conditions are imposed, the number of possible rules reduces considerably. In this paper we impose the conditions

of p-variate normality and equal covariance matrices for the K groups. Furthermore, we consider misclassification losses and prior probabilities of group membership to be the same across all groups.[1] Using the criterion of maximizing the (posterior) probability of correct classification, the above restrictions lead to a decision rule based on linear composites of the p variable scores (Morrison, 1976, p. 241; Rao, 1973, p. 575) called linear classification functions (LCFs). The sample LCFs are defined by

$$L_g = (\bar{y}_g' S^{-1})\, y - \tfrac{1}{2}\, \bar{y}_g' S^{-1} \bar{y}_g ,$$

where

$g = 1, \ldots, K,$

$\bar{y}_g$ = p-element vector of response variable means for group g,

S = (p x p) pooled covariance matrix for y-scores, and

y = p-element observation vector of y-scores.

It should be noted that there are K LCF scores obtainable for each unit, one for each group; that is, there are as many LCFs as there are groups. The decision rule is to assign a given unit to the group that yields the largest L_g value for that unit's observation vector.[2]

From the expression for the linear composite, L_g, it is seen that the p-element vector of weights is given by

$$b_g' = \bar{y}_g' S^{-1},$$

and the constant is $-\tfrac{1}{2}\, \bar{y}_g' S^{-1} \bar{y}_g$.

An approach to "interpreting" LCFs is to examine the within-groups correlations between the LCFs and the p response variables. [Calculations of these correlations yields p coefficients for each LCF; i.e., pK coefficients in total.] By so doing two types of information on each group are

obtainable: (a) A particular group can be characterized by high scores for variables with large positive correlations, and low scores for variables with large negative correlations,[3] and (b) the absolute values of the correlations yield a rank-ordering of the variables for contribution to classification accuracy for that particular group. That is, variables that best distinguish, in a classification accuracy sense, group g from the remaining groups can be identified. A few "high" LCF-variable correlations indicate that LCFs comprised of just those variables might be used for classification for that particular group, rather than the LCFs which involve all of the p variables.

It may be proven, however, that the ordering of the LCF-variable correlations for a given group is identical to that yielded by an ordering of means of standardized variable scores for the group. (See Appendix A for proof, including standardization procedure.) That is, the LCF-variable correlations for group g are proportional to the standardized variable means for group g.

This proof makes the first type of interpretive information very reasonable. That is, a given group of units may reasonably be characterized as a group having high (low) means on some variable(s) anc, possibly, low (high) means on some other variable(s). A typical member of a given group, then, may be considered to have a high (low) score on some variable(s) and, possibly, a low (high) score on some other variable(s). ["High" and "low" values are considered in terms of absolute standardized values.]

The proof also enables one to determine the relative contribution of variables to separate group classification accuracy without having to carry out elaborate calculations to obtain the within-groups LCF-variable correlations. It is only necessary to calculate the means of the standardized variable scores, a (pxl) vector which may be denoted as $\bar{x}_g$ for group g.

The calculation of the entries in the $\bar{x}_g$ vectors can be accomplished using matrix-manipulation statements through PROC MATRIX of the SAS computer package. (See a listing of statements in Appendix B.) [The LCF-variable correlations can be calculated using either (1) a two-step procedure of first determining the LCF weights and forming the K composites, followed by entering, for each unit, the K LCF scores and the p variable scores into a computer program (e.g., SPSS's DISCRIMINANT) to get the desired within-groups correlations, or (2) a computer program (see Appendix B) to obtain values in the next to last expression in the proof in Appendix A.]

Illustrative Data Sets

To illustrate the interpretation of linear classification functions (LCFs), two real data sets were analyzed. For the purposes of this paper, the research problem associated with each data set is assumed to be one of prediction of group membership. Data Set 1 involves three groups of college students, on each of whom were available 13 response variable measures. The multivariate normality condition is assumed met, and Box's test indicated that the condition of equal covariance matrices is tenable. The x-means for each group are given in Table 1. With this data set, the first group of

---

Insert Table 1 about here

---

individuals may be characterized as having low scores on 12 of the 13 variables, particularly variables 5, 11, and 12; the second group as having high scores on 3, 4, and 6; and the third group as having high scores on 5, 11, and 12. The characterization of groups 1 and 3 is, perhaps, more pronounced than that of group 2. [These characterizations may not be too surprising since, in some senses, the groups were considered "ordered" from low to high.]

In terms of classification accuracy, variables 5, 11, and 12 contribute most for groups 1 and 3, (i.e., these variables are most highly related with the respective LCFs), while variables 3, 4, and 6 contribute most for group 2. That is, variables 3, 4, and 6 best distinguish group 2 from groups 1 and 3. The classification results using all 13 variables are given in Table 2. If only variables 5, 11, and 12 are used, the separate group

---

Insert Table 2 about here

---

classification accuracy for groups 1 and 3 is pretty much upheld. For group 1 the number of "hits" drops from 30 to 26, and for group 3 from 32 to 30; this is a decrease for the two groups of only 8% when three variables are used rather than thirteen. [It should be noted that predictive accuracy will generally not decrease with a decrease in the number of variables, provided "good" variables are retained; the LCF-variable correlations indicate the good variables to retain.] Group 2 classification accuracy remains the same (71 hits) using only variables 3, 4, and 6 versus using all thirteen variables. From Table 1 it might be argued that variables 1 and 2 should also be considered in the group 2 characterization, and as important variables from a classification sense; in fact, if variables 1, 2, 3, 4, and 6 are used, the number of group 2 hits increases from 71 to 74.

Data Set 2 involves four groups of children, on each of whom were available 14 response variable measures. Again, the conditions of multivariate normality and equal covariance structure were considered tenable. The x-means for each group are given in Table 3. Group characterizations may be made as

---

Insert Table 3 about here

---

follows: group 1 is characterized by high scores on variables 2, 8, and 11, and low scores on variable 5; group 2 by low scores on variables 3, 8, 9, and 11; group 3 by high scores on variables 1, 9, and 13; and group 4 by low scores on variables 1 and 2, and high scores on variables 3 and 5. In terms of classification accuracy, variables 8 and 11 contribute most for group 1, variables 3, 8, 9, 11, and 14 for group 2, variables 1, 9, and 13 for group 3 and variables 1, 2, 3, and 5 for group 4. The fourteen variable classification results are given in Table 4. When only variables 8 and 11

Insert Table 4 about here

---

are used, the predictive accuracy for group 1 remains the same as when all fourteen variables are used. Group 2 accuracy only slightly decreases (from 8 "hits" to 7) when the "best" five variables instead of fourteen are used. Predictive accuracy for group 3 is low with all fourteen variables, and remains low for any subset; this is because group 3 is located "in-between" group 1 and 2, and not appreciably distant from either. For group 4, the predictive accuracy is affected very little when only variables 1, 2, 3, and 5 are used--14 hits resulted rather than 15.

By examining Table 3 it may be noted that variable 11 ranks high for groups 1 and 2 but low for group 4. This assessment of the importance of variable 11 is reflected by this variable alone yielding good classification results for group 1 (21 hits versus 16 using all fourteen variables) and group 2 (7 hits versus 8), but very poor results for group 4 (3 hits versus 15).

Summary

Techniques of discriminant analysis are appropriate for analyzing multivariate data based on two somewhat different research situations. The analysis in either situation calls for obtaining and interpreting linear composites of response variables. In the situation of group comparisons where a descriptive discriminant analysis is appropriate, the composites of interest are linear discriminant functions (LDFs). In the situation of group assignment where a predictive discriminant analysis is appropriate, the composites of interest are linear classification functions (LCFs). In general, LDFs and LCFs are determined differently, are different in number, are determined for different purposes, and are interpreted differently.

The focus of this paper is the interpretation of LCFs. LCFs may be "interpreted" according to LCF-variable within-groups correlations which are discussed herein. These correlations may be used for two purposes: (a) to characterize each group with respect to variable scores, and (b) to determine the relative contribution of the variables to separate group classification accuracy. It is proved that the LCF-variable correlations for a group are proportional to standardized variable means for that group. The proof enables a researcher to quite readily accomplish the two interpretation purposes.

Appendix A

Let Y be an N x p sample data matrix, with $Y' = [\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_N]$. The y-scores may be standardized into x-scores as follows:

$$X = (Y - \underline{1}\,\bar{\underline{y}}')\,D^{-\frac{1}{2}},$$

where

$\underline{1}$ = an N-element unit vector,

$\bar{\underline{y}}' = \frac{1}{N}\underline{1}'Y$, and

$D^{-\frac{1}{2}}$ = diagonal matrix of reciprocals of positive square roots of diagonal elements (s_{ii}, i = 1, ..., p) of S.

Then,

$$x_u^{(i)} = \frac{y_u^{(i)} - \bar{y}^{(i)}}{\sqrt{s_{ii}}} \quad \text{for variable i and unit u, and}$$

$$\bar{x}_g^{(i)} = \frac{1}{N_g}\sum_{u=1}^{N_g}\frac{y_{gu}^{(i)} - \bar{y}^{(i)}}{\sqrt{s_{11}}}, \quad g = 1, \ldots, K$$

$$= \frac{1}{\sqrt{s_{ii}}}\,(\bar{y}_g^{(i)} - \bar{y}^{(i)}).$$

Also,

$$\bar{\underline{x}}_g = (D^{-\frac{1}{2}})\,(\bar{\underline{y}}_g - \bar{\underline{y}}).$$

Therefore,

$$\bar{\underline{x}} = \frac{1}{N}\sum_{g=1}^{K} N_g\,\bar{\underline{x}}_g = \frac{1}{N}\,D^{-\frac{1}{2}}\sum_{g=1}^{K} N_g\,(\bar{\underline{y}}_g - \bar{\underline{y}})$$

$$= \underline{0}.$$

That is, the grand mean of the x-scores is zero for each variable.

Furthermore, the estimated variance of $\underline{x}'$ is

$$\begin{aligned} \text{est Var } (\underline{x}') &= \text{est Var } (\underline{y}' \, D^{-\frac{1}{2}}) \\ &= D^{-\frac{1}{2}} \, S \, D^{-\frac{1}{2}} \\ &= R, \end{aligned}$$

the (p x p) within-groups correlation matrix for Y. The LCFs in standard-score form are given by

$$\begin{aligned} L_g^* &= (\bar{\underline{x}}_g' \, R^{-1}) \, \underline{x} - \tfrac{1}{2} \, \bar{\underline{x}}_g' \, R^{-1} \bar{\underline{x}}_g \\ &= \underline{b}_g^{*\prime} \underline{x} + c_g. \end{aligned}$$

Now,

$$\begin{aligned} \text{est Var } (L_g^*) &= \text{est Var } (\bar{\underline{x}}_g' \, R^{-1} \, \underline{x}) \\ &= \bar{\underline{x}}_g' \, R^{-1} (R) \, R^{-1} \bar{\underline{x}}_g \\ &= \bar{\underline{x}}_g' \, R^{-1} \, \bar{\underline{x}}_g. \end{aligned}$$

But,

$$\text{est Var } (x^{(i)}) = r_{ii} = 1.$$

And,

$$\begin{aligned} \text{est Cov } (L_g^*, \underline{x}') &= \bar{\underline{x}}_g' \, R^{-1} \, R \\ &= \bar{\underline{x}}_g'. \end{aligned}$$

Hence,

$$\begin{aligned} \text{est Corr } (L_g^*, x^{(i)}) &= \frac{\bar{x}_g^{(i)}}{\sqrt{(\bar{\underline{x}}_g' \, R^{-1} \, \bar{\underline{x}}_g) \, 1}} \\ &= \frac{\bar{x}_g^{(i)}}{\sqrt{\underline{b}_g^{*\prime} \, \bar{\underline{x}}_g}}. \end{aligned}$$

Therefore, $\text{est Corr } (L_g^*, x^{(i)}) \propto \bar{x}_g^{(i)}$.

APPENDIX H

```
*                CALCULATION OF STANDARDIZED MEANS                    ;
*                  AND LCF-VARIABLE CORRELATIONS                      ;


PROC MATRIX;
FETCH X DATA=TR;                          *X,X1-X4, DATA MATRICES FOR   ;
FETCH X1 DATA=GP1; K=1;                   *  THE OVERALL SAMPLE AND FOR ;
FETCH X2 DATA=GP2; K=K+1;                 *  EACH GROUP.IN THIS EXAMPLE ;
FETCH X3 DATA=GP3; K=K+1;                 *  K=4                        ;
FETCH X4 DATA=GP4; K=K+1;
NVAR=NCOL(X);                             *NO. COLUMNS IS NO. OF VARS.  ;
NOBS=NROW(X);                             *NO. ROWS IS NO. OF OBS.      ;
NOBS1=NROW(X1);
NOBS2=NROW(X2);
NOBS3=NROW(X3);
NOBS4=NROW(X4);
JV=J(1,NOBS);                             *CREATE UNIT VECTORS          ;
JV1=J(1,NOBS1);
JV2=J(1,NOBS2);
JV3=J(1,NOBS3);
JV4=J(1,NOBS4);
JVV=J(K,1);
YBARP=(JV*X)#/NOBS;                       * CALCULATE K+1 VECTORS OF    ;
YBAR1=(JV1*X1)#/NOBS1;                    *  GROUP MEANS                ;
YBAR2=(JV2*X2)#/NOBS2;
YBAR3=(JV3*X3)#/NOBS3;
YBAR4=(JV4*X4)#/NOBS4;
YBARGP=YBAR1//YBAR2//YBAR3//YBAR4;        *CONCATENATE VECTORS OF MEANS ;
YBARG=YBARGP';
SUM1=JV1*X1;                              *CALCULATE SUMS OF SQUARES AND;
XPX1=X1'*X1;                              *  CROSS PRODUCTS             ;
XPX1=XPX1-(SUM1'*SUM1)#/NOBS1;
SUM2=JV2*X2;
XPX2=X2'*X2;
XPX2=XPX2-(SUM2'*SUM2)#/NOBS2;
SUM3=JV3*X3;
XPX3=X3'*X3;
XPX3=XPX3-(SUM3'*SUM3)#/NOBS3;
SUM4=JV4*X4;
XPX4=X4'*X4;
XPX4=XPX4-(SUM4'*SUM4)#/NOBS4;
XPX=XPX1+XPX2+XPX3+XPX4;
VCV=XPX#/(NOBS-K);                        *FIND WITHIN-GROUP COVARIANCE ;
                                          *  MATRIX                     ;
SQT=SQRT(VCV);                            *OBTAIN DIAGONAL MATRIX OF    ;
RC=RECIP(SQT);                            *  RECIPROCALS OF STD. DEVS.  ;
D=DIAG(RC);
R=D*VCV*D;                                *CALCULATE CORRELATION MATRIX ;
RI=INV(R);
YBARPP=JVV*YBARP;                         *FORM MATRIX OF GRAND MEANS   ;
YBAR=YBARPP';
DIFF=YBARG-YBAR;                          *CALCULATE MATRIX OF STANDARD-;
XBAR=DIFF'*D;                             *  IZED MEANS (SEE APPENDIX A);
MEANS=XBAR';
XPX=MEANS'*RI*MEANS;                      *OBTAIN LCF-VARIABLE CORRELA- ;
SXPX=SQRT(XPX);                           *  TIONS (SEE APPENDIX A)     ;
DXPX=RECIP(SXPX);
DXPX=DIAG(DXPX);
COR=MEANS*DXPX;
PRINT MEANS COR;                          *PRINT MEANS AND CORRELATIONS ;
```

Reference Notes

1. Bray, J. H., & Maxwell, S.E. Analyzing and interpreting significant MANOVAS. Paper presented at the Annual Meeting of the American Psychological Association, New York, September 1979.

2. Bavry, J. A simulation study of MANOVA canonical variate interpretation . Paper presented at the Annual Meeting of the American Educational Research Association, Toronto, March 1978.

References

Borgen, F.H., & Seling, M.J. Uses of discriminant analysis following MANOVA: Multivariate statistics for multivariate purposes. Journal of Applied Psychology, 1978, 63, 689-697.

Green, B.F. The two kinds of linear discriminant functions and their relationship. Journal of Educational Statistics, 1979, 4, 247-260.

Huberty, C.J. Discriminant analysis. Review of Educational Research, 1975, 45, 543-598.

Morrison, D.F. Multivariate statistical methods (2nd ed.). New York: McGraw-Hill, 1976.

Nie, H.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., & Bent, D.H. Statistical package for the social sciences (2nd ed.). New York: McGraw-hill, 1975.

O'Grady, K.E. Comments on "Some statistical methods for the assessment of multiple outcome criteria in behavioral research." Behavioral Therapy, 1978, 9, 471-473.

Rao, C.R. Linear statistical inference and its applications (2nd ed.). New York: Wiley, 1973.

Sanathanan, L. Discriminant analysis. In D.J. Amick & W.J. Walberg (Eds.), Introductory multivariate analysis. Berkeley, Calif.: McCutchan, 1975.

Spector, P.E. What to do with significant multivariate effects in multivariate analysis of variance. Journal of Applied Psychology, 1977, 62, 158-163.

Tatsuoka, M.M. Multivariate analysis. New York: Wiley, 1971.

Tatsuoka, M.M. Classification procedures. In D.J. Amick & H.J. Walberg (Eds.), Introductory multivariate analysis. Berkeley, Calif.: McCutchan, 1975.

Weiss, D.J. Multivariate procedures. In M.D. Dunnette (Ed.), Handbook of industrial and organizational psychology. Chicago: Rand McNally, 1976.

Wilkinson, L. Confirmatory rotation of MANOVA canonical variates. Multivariate Behavioral Research, 1977, 12, 487-494.

Footnotes

[1]For constant losses and priors, and normality, but for unequal covariance matrices, what are used is quadratic classification functions (QCFs). The sample QCFs are define by

$$Q_g = -\tfrac{1}{2} \ln|S_g| - \tfrac{1}{2} (\underline{y} - \bar{\underline{y}}_g)' S_g^{-1} (\underline{y} - \bar{\underline{y}}_g),$$

where S_g is the covariance matrix for group g.

[2]If LDF scores are used as input, then the $\bar{y}$ vector would have at most r elements, where r = min (p, K - 1).

[3]In making such characterizations one should consider pair-wise separation among the K groups. This involves determining K(K-1)/2 D^2 values, where

$$D^2_{gg'} = (\bar{\underline{y}}_g - \bar{\underline{y}}_{g'})' S^{-1} (\bar{\underline{y}}_g - \bar{\underline{y}}_{g'}).$$

If a D^2 value is essentially zero, then group characterizations might be based on LCF-variable correlations involving K-1 groups.

Table 1

The x-Means for Data Set 1

Variable	Group 1	2	3
1	-0.61	0.15	0.24
2	-0.28	0.15	-0.07
3	-0.61	0.26	0.00
4	-0.49	0.21	0.01
5	-1.10	0.07	0.88
6	-0.60	0.21	0.12
7	-0.52	0.01	0.47
8	-0.14	0.07	-0.03
9	-0.37	0.04	0.26
10	-0.30	-0.01	0.31
11	-1.31	-0.08	1.42
12	-1.78	0.02	1.65
13	0.42	0.05	-0.51

Table 2

Classification Results for Data Set 1

		Predicted Group			
		1	2	3	Total
Actual Group	1	30	5	0	35
	2	7	71	3	81
	3	0	5	32	37
					153

Table 3

The x-Means for Data Set 2

	Group			
Variable	1	2	3	4
1	0.11	-0.15	0.34	-0.26
2	0.25	-0.11	0.11	-0.28
3	-0.11	-0.45	0.13	0.34
4	0.11	-0.25	-0.15	0.16
5	-0.24	0.11	-0.08	0.26
6	0.12	-0.27	0.17	-0.08
7	0.12	-0.36	0.17	-0.02
8	0.36	-0.49	0.11	-0.16
9	0.13	-0.42	0.29	-0.07
10	-0.01	-0.04	0.01	0.03
11	0.45	-0.75	-0.05	0.05
12	0.12	-0.01	0.06	-0.18
13	0.17	-0.25	0.26	-0.21
14	0.22	-0.41	0.00	0.03

Table 4

Classification Results for Data Set 2

		Predicted Group				
		1	2	3	4	Total
Actual Group	1	16	2	3	4	25
	2	4	8	1	2	15
	3	6	3	2	5	16
	4	3	3	1	15	22
						78